GSSCC DNA Special Interest Group handout - - March 3, 2026

Endogamy, Multiple Relationships and Pedigree Collapse

compiled by Gail Burk

Let's examine three situations that can affect analysis of our DNA matches. Diahan Southard recommends that we consider two factors in dealing with unexpected instances of increased DNA sharing - - how many of these types of events are there in your family tree, and how far back in time did these events take place.

Endogamy - - When people in a group (religious, cultural, geographical, ethnic or tribal) partner within that group for generations. There will be numerous instances of cousins marrying cousins; it will have gone on for generations; and it may extend far back in time. That is, numerous instances, and extending far back in your family tree.

In endogamous populations, everyone will descend from the same gene pool. People will be related to each other within a recent genealogical time frame on multiple ancestral pathways and the same ancestors will, therefore, appear in many different places on their pedigree chart. Endogamy can be the result of a conscious decision or cultural pressure to marry within the selected group, but also occurs as a result of geographical isolation (for example, in island communities).

Examples of endogamous groups include Ashkenazi Jews, Polynesians, Low German Mennonites, the Amish, Acadians, French Canadians, people from many Arab countries, people from Newfoundland, and people from many islands. Endogamy is also a problem in some early Colonial American populations.

Analyzing DNA results from endogamous populations can be especially challenging because people will routinely have large numbers of matches in the DNA databases. Basically, you will share too much DNA with all of your cousins, to accurately reflect their actual genealogical relationships. You will share DNA with others in your endogamous community not because you share a single recent common ancestor, but because you share numerous more distant common ancestors. Interpretation of autosomal DNA matches can be particularly difficult, especially in the case of endogamous populations where the pedigrees cannot be traced back beyond the 1800s.

One way you can tell the difference between someone who is closely related and someone who shares DNA because of endogamy is to pay attention not just to the total amount of DNA shared, or your relationship prediction, but to look at the size of the segments of shared DNA. Diahan Southard recommends that we prioritize matches who share a longest segment of at least 20 cM. For people with Jewish ancestry it has been suggested that you need to have at least one long segment of 23 cM or more in order to have some hope of finding the relationship in genealogical time.

Diahan also recommends using adjusted cM amounts when evaluating your endogamous DNA matches. At Ancestry, this means using the Timber-adjusted cM figure rather than the unweighted DNA figure. You want to prioritize matches that have at least one segment of 20 cM or more, and matches in which the gap between the Timber cM amount and the unweighted cM amount is small. At companies with chromosome browsers, Diahan recommends looking at all the shared segments and eliminating from consideration any segment that is 10 cM in length or smaller. If you have several small segments, add up the cMs for those segments and subtract that number from the shared cM to arrive at an adjusted amount of shared DNA. Making the adjustment will assist you in calculating the category of cousin relationship you share with your endogamous DNA matches.

When relationships are more distant than predicted, coefficients of relationship can be used to compute the expected quantity of shared cM. The term coefficient of relationship is used by geneticists to quantify the expected amount of shared DNA between related individuals. The important elements of the coefficient of relationship are (1) identifying the distinct inheritance pathways that exist between the people of interest; and (2) determining the degrees of relationship that separate them along those distinct pathways. There is math involved! Read about coefficients of relationship here:

https://isogg.org/wiki/Coefficient_of_relationship

Multiple Relationships - - Multiple relationships occur when you are related to someone on more than one of your family lines, thus you have multiple ancestral couples in common; e.g., if cousins married cousins. When you and your DNA match share more than one ancestor or ancestral couple as a 2x great-grandparents or closer, then you have a situation of multiple relationships. Key to this definition is the question of how many instances, and how close in time it is. If it is just one instance of sharing more than one ancestor with your match, and the time is close - - at 2x great-grandparent or closer - - you are dealing with multiple relationships. Multiple relationships that are farther back generationally in your tree will have less effect of your shared amounts of DNA

If you have a recent multiple relationship situation, you will likely see the amount of DNA for each relationship reflected in your total cM count. Diahan Southard recommends that you consult the Shared Centimorgan Project at DNA Painter to estimate how much shared DNA on average you would have for each relationship, then add those two numbers together.

An important step in analyzing DNA matches is to view the histogram for the proposed relationships at the Shared cM Project. The histogram can help you see where your elevated amount of shared DNA fits on the bell curve, and what percentage of DNA testers are represented at that spot. In matches without multiple relationships, you'd expect that some matches will fall below the average shared DNA, and some higher. When your matches are consistently on the high end of shared DNA, it can signal multiple relationships. You can access the Shared Centimorgan Project here:

https://dnapainter.com/tools/sharedcmv4

The problem is, if you don't know that you have multiple relationships, you'll try to fit your DNA match into a closer relationship (based on the elevated shared cMs or predicted relationship), and you won't understand why that doesn't work. Sometimes, the total amount of shared DNA is far enough outside of the expected range to alert you to the problem. Multiple relationships may not affect the DNA so much that it changes the relationship estimate, however, especially if it's at the third cousin level or more distant. If someone shares more than two multiple relationships, however, it may have a greater impact on the shared DNA.

Clustering your matches is a tool used to help discover the common ancestor between a test-taker and a DNA match. But clustering can be confusing in the case of multiple relationships, when you look at shared matches and they fall into more than one cluster. Clustering can be impossible altogether with endogamy, where all your matches form a single, huge messy cluster. When you see these anomalies in your clusters, that's a signal to do genealogy to research your match's ancestors. You can use network graphs to visualize familial relationships and cluster matches based on shared DNA. In the case of multiple relationships and endogamy, network graphs can help you visualize the inter-relatedness of your matches. Learn more about network graphs here:

https://myfamilypattern.com/network-graphs/

Here are some of Diahan Southard's take-aways and tips for multiple relationships - - Be careful not to confuse multiple relationships with pedigree collapse; build complete trees for DNA matches to check for multiple shared ancestors; evaluate each DNA match with the Shared cM Project Tool and explore inflated amounts of DNA. The amount of shared DNA will be more obviously affected by recent generations than by distant ones.

Pedigree Collapse - - In genealogy, pedigree collapse occurs when two individuals knowingly or unknowingly share an ancestor. This causes the family tree of their offspring to be smaller than it would otherwise be. For example, the offspring of two first cousins has at most only six great-grandparents instead of the normal eight. Again, it is key to determine how close in time, and how many instances have occurred.

In some cultures, cousins were encouraged or required to marry in order to maintain kinship bonds or retain wealth and property within a family. Among royalty, the frequent requirement to only marry other royals resulted in a reduced gene pool in which most individuals were the result of extensive pedigree collapse. More generally, intermarriage may frequently occur within a small village, with a limited available gene pool.

Most historians consider the House of Habsburg as an example of genetically-induced disease as the direct result of pedigree collapse. The lineage of the last Habsburg King of Spain, Charles II, was so intermarried that he had a higher inbreeding ratio than if he had been born to a brother-sister couple. The maximum degree of pedigree collapse of 50% within a single generation is caused by procreation between full siblings. Such children have only two different grandparents, instead of four.

Small, isolated populations such as those on remote islands (e.g., Easter Island and Iceland) represent extreme examples of pedigree collapse, but the common historical tendency to marry those within walking distance, due to the relative immobility of the population before modern transport, meant that many marriage partners were at least distantly related. Even in America as late as the 19th century, immigrants tended to marry among their own ethnic, language or cultural group, resulting in many cousin marriages. Some geneticists believe that everybody on Earth is at least 50th cousin to everybody else.

At some point in everyone’s pedigree, instead of different individuals filling out your chain of generations, some of the same ancestors will start to appear multiple times. The term pedigree collapse aptly describes this situation, as when repeat ancestors fill in pedigree slots the inverted triangle starts to collapse in on itself and look more like a diamond. Although this scenario plays out in everyone’s pedigree at some point, it becomes relevant to autosomal DNA test takers when there is a degree of pedigree collapse in the most recent several generations. This comes into play the most in your match list, where you will see elevated levels of DNA sharing between cousins that descend through the same ancestors that are involved in the pedigree collapse.

The quantity of DNA you share with your matches may not be a stand-alone clue that pedigree collapse is at work. It depends on how many instances there are, and how close in time they are to you and your match. An isolated incident of recent pedigree collapse may, or may not, be obvious in the cM count you share with your cousin; the cM amount might still fall within the expected range for their expected degree of relationship. In pedigrees both with and without pedigree collapse, there is an expected average level of DNA, but that average is in the middle of a range of shared DNA for the expected relationship. If you and your match are still within that range, you won’t necessarily be tipped off to a situation of pedigree collapse.

As Diahan Southard explains, on average, children inherit 50% of their DNA each from their mom and dad, 25% from each of their 4 grandparents, 12.5% from each of their 8 grandparents, 6.25% from each of their 16 great-grandparents, and so forth. The average amount of DNA inherited from an individual ancestor is halved going back each generation level beginning with the parents. It's important to note that these are just averages and the actual amount of DNA inherited is random but will fall within a range around the average numbers.

Further developing this idea, cousins inherit a predictable amount of DNA from their ancestor, but they don't necessarily inherit the same segments of DNA. For instance, first cousins will inherit on average 25% of DNA from one of their common grandparents, but it is expected that on average they will only inherit 6.25% of the same DNA (that's 25% of the 25%).

It's also important to note that, due to random inheritance, the farther back the ancestor, the less likely that there are segments large enough in current descendants today to be detected by the types of markers and methods currently used by ancestry testing companies. For instance, AncestryDNA estimates that only 32% of documented fifth cousins will share enough DNA inherited from their common ancestor for that relationship to be detected.

To calculate shared DNA in pedigree collapse, Diahan Southard suggests applying these considerations: If you see a match where you both descend through ancestors involved in pedigree collapse, what degree of elevated DNA sharing would you expect to see? Would it be enough to bump you up into a different relationship category? For example, if you were second cousins, would the elevated shared DNA be enough to make you look like you were first cousins? Matches with pedigree collapse might be first cousins by way of common grandparents; but if the grandparents were also first cousins to each other; the current DNA matches would be more genetically similar than cousins who descend through unrelated grandparents. This causes an elevated level of shared DNA in downstream descendants over what they would expect in the absence of pedigree collapse.

An isolated incident of pedigree collapse, however, is likely not enough to inflate their amount of shared DNA to make them appear to belong to a different relationship category. The effect of the pedigree collapse is lessened with each succeeding generation, as descendants become more distant from the "double-dose" ancestors that appear twice in their pedigree. More distantly related unions such as second cousin, etc., would have even less of an effect on the total shared DNA between downstream descendants. Cousins resulting from these more distantly related unions who encounter each other for the first time on a DNA company match list would likely see no evidence of this more distant and isolated pedigree collapse in their amount of shared DNA.

Diahan's take-away is that pedigree collapse always results in elevated levels of DNA sharing in the next several generations of descendants, but the farther away in time from the first cousin relationship, the less effect it has on the shared DNA of descendants. If the presence of pedigree collapse is isolated rather than widespread, downstream cousins will still expect increased DNA sharing but the elevated amount is modest enough to keep the augmented values well within the range of expected shared DNA for that relationship. Cousins will likely not recognize from their amount of shared DNA alone that they have an isolated incident of pedigree collapse in their recent pedigree. This changes, however, if the incidents of closely related unions are widespread, as the augmented level of cM sharing among descendants may be enough to bump them up into a closer cousin category. In the presence of widespread pedigree collapse, the same tools used to determine the coefficient of relationship among relatives can be applied.

Founder Effect - - It's appropriate to also mention founder effect. Although founder effect and endogamy are different, they can occur concurrently in your family tree. Founder effect is a principle from population genetics that describes the reduced diversity seen in a group of descendants who come from an isolated group that was established by just a few people. Both founder effect and endogamy can lead to reduced diversity in your DNA. Founder effect can make the people from this descendant cohort genetically distinct in a way that is detectable by DNA analysis.

The founder effect is seen in historical populations, like some of our ancestors' communities. One example of the possibility of founder effect can be seen in early American colonists. It is possible to detect a general "Puritan genetic community" for the New England region, as this population separated and continued to generally isolate themselves for generations. Many times, however, the founder effect on descendants of founding population groups is not profound enough to leave a distinct genetic legacy.

Founder effect and endogamy are not the same thing. A founder effect may result in reduced genetic diversity in a community. But for endogamy to come into play with your DNA analysis, that community would have to remain genetically isolated for a long time. Both can operate concurrently in small isolated populations, however, and contribute to elevated DNA sharing in descendants. It takes a lot of time for the genetic signature of endogamy to develop - - hundreds or thousands of years of extended endogamous practice.

Conclusion - - Look at your own family tree. Do you have instances of these three situations in your ancestral lines? All of us are likely to see at least one of these occurrences among our ancestors. Some of us will hit the jackpot, and see instances of endogamy, multiple relationships, and pedigree collapse in our family trees. And maybe we will get a bonus prize for also having instances of founder effect. As Diahan Southard constantly reminds us, one of the best ways to sort everything out is to do genealogy!

Sources

Your DNA Guide blog, by Diahan Southard
https://www.yourdnaguide.com/ydgblog/what-is-endogamy

Your DNA Guide blog, Part 1, by Jayne Elkins
https://www.yourdnaguide.com/ydgblog/pedigree-collapse-and-genetic-relationships
https://www.yourdnaguide.com/ydgblog/calculating-the-pedigree-collapse-effect-in-your-dna-matches
https://www.yourdnaguide.com/ydgblog/founder-effect

International Society of Genetic Genealogy
https://isogg.org/wiki/Endogamy
https://isogg.org/wiki/Pedigree_collapse

Family Locket blog, by Diana Elder
https://familylocket.com/endogamy-pedigree-collapse-and-multiple-relationships-whats-the-difference-and-why-does-it-matter/